# PFRA'S WESTERN GRAIN TRANSPORTATION PAYMENT PROGRAM LANDCOVER GENERALIZATION PROCESS

**James T. Ashton**
Agriculture Agri-Food Canada, Prairie Farm Rehabilitation Administration
ashtonj@em.agr.ca

**Abstract**
*One of the greatest research problems in GIS is the fascinating phenomenon of map generalization, where maps lose information dependant on their source scale. Generalization effects are an inherent characteristic of all geographic data and since they can introduce serious errors and aerial misrepresentations it is imperative to know what these effects are and how large they can be. Raster-based generalization techniques have received less attention than the vector-based techniques. Yet for area data, it would seem to be intrinsically suitable since the basic raster unit, the grid cell, is itself an area. This paper discusses some of the initial difficulties and issues addressed in deriving an algorithm for PFRA's automated generalization process for the classified imagery gathered during the Western Grain Transportation Payment Program (WGTPP) 1993-1995.*

## Introduction
The Landcover Generalization process was undertaken to solve rendering problems of the original vectorized landcover data due to its unwieldy/overwhelming size. LANDSAT 7 imagery used in the process was collected during the WGTPP. This landcover imagery has a 30 meter resolution and is stored in over 1,100 vectorized 1:50,000 map sheet tiles. The data requires over seven gigabytes of disc space. If the user wishes to view a specific area, they must identify the 1:50,000 map sheet, then browse to the desired mapsheet. Due to the great detail, this vector product (l02 in GIS_LIB) becomes computationally impossible on the desktop computer when multiple mapsheets are drawn. Spatial analysis and modeling at provincial or western Canada extents are also impossible in the mapsheet format.

The WGTPP Generalized Landcover is a seamless vector coverage, covering the agricultural extent of Western. The generalized vector coverage is 124 MB in size, and requires roughly 60 times less space than the mapsheet tiles, thus making it easy to be used with desktop GIS applications. An algorithm was created using the processing tools of the ArcINFO GRID module. A series of AMLs and digital rules have been implemented in the event that new satellite imagery is acquired.

## Objective
The essential objective of the generalization is to preserve the characteristics and integrity of geographic data while reducing the level of detail in its representation.

There were 6 goals for the Generalized Landcover:
- To be used cartographically with outputs at a scale of 1:1 million.
- As a viewing and analytical tool within PFRA regions and districts
- A planning tool, and a 'Quick' statistical tool for analytical staff
- To be incorporated into Internet Mapping applications
- To develop and establish a process with digital rules of generalization for use with new satellite imagery

- Collaborating in partnership with various levels of government, the private sector and the academic community to capitalize on their collective expertise and to ensure seamless public delivery of geospatial information.

**Methods**

The generalization process began with four 30 meter grids which were split by province. Since each provincial GRID possessed over 20,000 rows and columns, they were as much as the processing power of the ArcINFO software could manage. Each GRID was projected into a common working projection of Lamberts Conic Conformal.

Another initial decision or digital rule implemented was to select all landcover classes that were Unclassified or Clouds/Shadow and replace them with their nearest neighbor. The rational behind this decision is, these areas would likely be lost in the generalization and the removal of any known erroneous data would potentially create a better representation of the specific area.

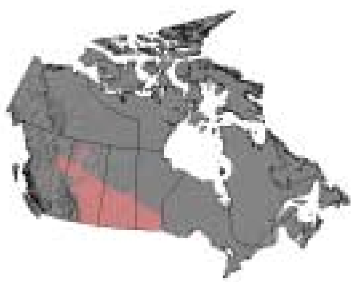
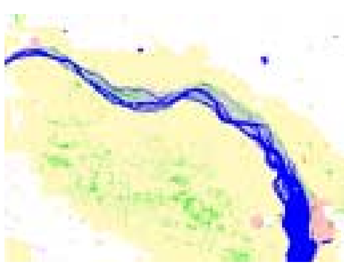


***Figure 1.*** This map shows the extent of coverage of the lancover data.

Initial attempts with the generalization process for the 30 meter imagery proved to be challenging. As one might expect rare landcover types were lost when resolution became coarser (Turner et al, 1989). This rate of loss is closely intertwined with the spatial pattern of the distribution of landcover types. Clumped rare landcover types disappeared slowly or were retained with increasing grain but dispersed landcover types were lost rapidly. These initial attempts of generalization of Blockmajority and Focalmajority filters lead to unacceptable results. In Saskatchewan for example, the domination of the cropland class proved as expected, to replace smaller more dispersed classes such as wetlands and shrubs. Furthermore, initial attempts of generalizing British Columbia were completely different than in Saskatchewan, Alberta or Manitoba. The dominant and dispersed nature of the landcover class of Trees, made BC's situation fundamentally different. Other errors such as connectivity were induced with basic filters. The detail of the original imagery, depicting such features as a braided stream became to complex for a simple automated generalization operation. Attempting to mimic such a subjective and intuitive procedure, the complex diverse and non-deterministic automation results were unacceptable.

***Figure 2a and 2b,*** The below illustrates initial difficulties of connectivity in complex riparian areas.

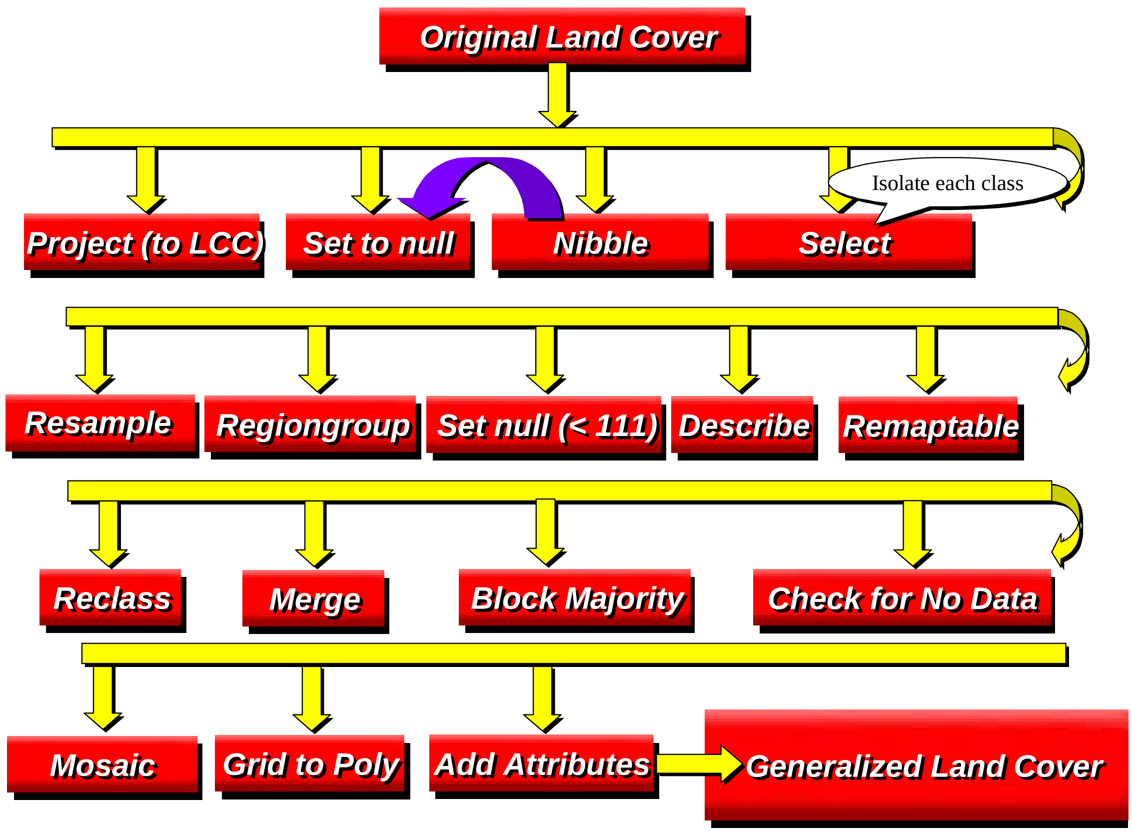
***2a.*** Original



***2b.*** Generalized

With the initial exploration in raster-based generalization, and the familiarity with the diverse and complex nature of the original imagery it became apparent that working with the landcover classes individually may be a viable solution. Figure 3. illustrates each step of the PRFA generalization process.

*Figure 3.* **Algorithm for PFRA's Landcover Generalization process.**




***1 . Original Landcover*** – Begin with the extent of data provincially. (ie Saskatchewan, Alberta, Manitoba, or Northern British Columbia).

***2. Project to LCC*** – Project the GRID to Lamberts Conic Conformal.

***3. Set to null*** – Set landcover classes Clouds/Shadow and Unclassified to null or (NODATA)

***4. Nibble*** – The resulting NODATA cells are replaced by the nearest neighbor.

***5. Select*** – Each landcover class is selected out to be processed individually (until the merge generalization operator Step 12).

***6. Resample*** – Each individual landcover class is resampled from 30 meter cell to 100 meter cell using nearest neighbor.

***7. Regiongroup*** – The resulting cells are then grouped by connectivity of individual landcover classes.

***8. Set null (-111)*** – groups with a count of less than 111 cells (ie. Less than1 sq KM), are removed from further processing. This is a minimum mapping unit digital rule which is implemented for this stage. However, after the landcover classes are merged back to a single raster this digital rule is not enforced. Enforcing this rule after the merge or before final raster to vector conversion in effect erased the significance of ranking the merge process by spatial distribution (see step 12).

***9. Describe*** – This raster is then described, the characteristics of the raster are noted to create the remap table.

***10. Remaptable*** – A remap table is created based on the characteristic of the describe.

***11. Reclass*** – Reclass each cell to its original landcover class. This process is needed as the value representing the landcover class has been altered in the value attribute table (.vat) of the GRID. The .vat is similar to the polygon attribute table (,pat) but it just has a number representing a class. In order for classes to be distinguished in the merge, the must all have the same value in the .vat ie Waterbodies is represented by 6 in the ,vat.

***12. Merge*** - All individual classes are merged back into one seamless provincial raster. Rasters are merged in an order of precedence. For example, when merging each individual class overlapping area is determined by the order that the class is entered in the argument list of the merge command. Within each province the distribution of wetlands was by far the most dispersed and had the smallest average patch size, it was regarded as highest priority. Each province again had different internal distribution and thus, had a different order of merging. Since the issue of connectivity of Waterbodies was important, it was generally second in the argument list.

***13. Blockmajority*** - An aggregate function that partitions the input grid into blocks finds the majority value (the value that appears most often) for the specified cells (defined by the neighborhood parameters) within the blocks, and sends it to the cell locations in the corresponding blocks on the output grid. In investigation and experience with the PFRA Landcover Generalization several different sized filters were used on different landcover classes. The parameters for each command used in the process are not included in this paper. The methods and parameters used for this process produced the best results of generalization with the original dataset. Duplicating the process with different spatial patterns and distributions of landcover types would likely warrant different sized filters.

***14. Check for nodata*** – Nodata slivers were replaced with their nearest neighbor.

***15. Mosaic*** – The merged provincial grids were mosaiced together to form a seamless agricultural extent raster.

***16. Gridpoly*** - The resulting generalized raster was vectorized into an ArcINFO coverage.

***17. Add Attributes*** – Two attributes, 'class' and 'acres' were added to the polygon attribute table (pat). Class, is the text description of the landcover class, ie Cropland. Acres is a calculation of the Area field in the .pat.

## Results

The actual reported cropland acres from the WGTPP, the 1996 Statistics Canada census numbers and the original imagery were used to verify accuracy.

Figure 5, depicts a Generalized Landcover cropland class that is remarkably close to the official data that was collected and verified during the WGTPP. It has been suggested that the percentage of some landcover types in the generalized format (such as Cropland) may actually be more accurate than the original imagery. For example, at the time the image was taken many areas may have been classed as a slough, or wetland only to be put into crop when suitable.

***Figure 5.*** The statistics for the Generalized Landcover are remarkably close to the reported information compiled from the Western Grains Transition Payment Program (WGTPP)

| | 96' CENSUS DATA | WGTPP PAID ACRES | ORIGINAL LANDCOVER | GENERALIZED LANDCOVER |
|---|---|---|---|---|
| **Manitoba** | 10,560,014 | 10,497,056 | 10,951,200 | 10,825,498 |
| **Sask** | 43,838,496 | 44,083,262 | 43,790,730 | 44,618,333 |
| **Alberta** | 22,387,133 | 22,032,142 | 21,322,370 | 22,792,180 |
| **BC** | 634,526 | NA | 544,294 | 488,477 |

On a provincial or larger extent the landcover class area distributions are within 4% of the original 30 meter imagery. For example, in the province of Saskatchewan the original imagery contained 56% cropland and the generalized vector coverage contains 58 % cropland.

## Conclusions

As with manual generalization, automated generalization is taken as a process of selecting and simplifying a description of geographic phenomenon. However unlike manual generalization that involves the simultaneous application of different factors and intuitiveness by a user in order to automate generalization it is necessary to break the process up into a series of smaller, identifiable steps, and eventually into an algorithm. (Joao, 1998). In conclusion, the PRFA Generalized Landcover now offers the user easy access to a value-added remotely sensed data. This data is accurate enough for use with provincial or western Canadian extent applications and modeling, and it is manageable on the desktop computer. Also there is a process and set of digital rules in the event new satellite imagery is acquired for the area.

**References**

Turner, M.G., O'Neill, R.V. Gardner, R.H., and Miline, B.T. 1989 Effects of changing Scale on the analysis of landscape pattern. Landscape Ecology. 3 / 4 (3):153

ESRI, 2000, "Map Generalization in GIS: Practical Solutions With Workstation ArcINFO Software". Technical Paper. It can be found in the White Papers section of Arconline at http://arconline.esri.com

Joao, E.M., 1998. "Causes and Consequences of Map Generalisation". Taylor & Francis.